Diversity dynamics of Phanerozoic terrestrial tetrapods at the local community scale

Close, Roger; Benson, Roger B J; Alroy, John; Behrensmeyer, Anna; Benito Moreno, Juan; Carrano, Matthew T.; Cleary, Terri; Dunne, Emma; Mannion, Philip D.; Uhen, Mark; Butler, Richard

DOI:
10.1038/s41559-019-0811-8

Diversity dynamics of Phanerozoic terrestrial tetrapods at the local-community scale

Roger A. Close[1], Roger B. J. Benson[2], John Alroy[3], Anna K. Behrensmeyer[4], Juan Benito[1,5], Matthew T. Carrano[4], Terri J. Cleary[1,6], Emma M. Dunne[1], Philip D. Mannion[7], Mark D. Uhen[8], Richard J. Butler[1]

[1]School of Geography, Earth and Environmental Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK.

[2]Department of Earth Sciences, University of Oxford, Oxford OX1 3AN, UK.

[3]Department of Biological Sciences, Macquarie University, NSW 2109, Australia.

[4]Department of Paleobiology, National Museum of Natural History, Smithsonian Institution, Washington DC 20013, USA.

[5]Department of Biology and Biochemistry, Faculty of Science, University of Bath, Claverton Down, Bath BA2 7AY, UK.

[6]Department of Earth Sciences, Natural History Museum, Cromwell Road, London SW7 5BD, UK.

[7]Department of Earth Science and Engineering, Imperial College London, London SW6 2AZ, UK.

[8]Department of Atmospheric, Oceanic, and Earth Sciences, George Mason University, Fairfax, VA 22030, USA.

The fossil record provides one of the strongest tests of the hypothesis that diversity within local communities is constrained over geological timescales. Constraints to diversity are particularly controversial in modern terrestrial ecosystems, yet long-term patterns are poorly understood. Here we document patterns of local richness in Phanerozoic terrestrial tetrapods using a global dataset comprising 145,332 taxon occurrences from 27,531 collections, much larger than that used any previous study. We show that local richness of non-flying terrestrial tetrapods has risen asymptotically since their initial colonization of land, increasing at most three-fold over the last 300 million years. Statistical comparisons support phase-shift models, with most increases in local richness occurring: 1) during the colonization of land by vertebrates, concluding by the late Carboniferous, and 2) across the Cretaceous/Palaeogene boundary. Individual groups such as mammals, lepidosaurs and dinosaurs also experienced early increases followed by periods of stasis often lasting tens of millions of years. Mammal local richness abruptly tripled across the Cretaceous/Palaeogene boundary, but did not increase over the next 66 million years. These patterns are consistent with the hypothesis that diversity is constrained at local community scales.

There is substantial disagreement about how the exceptional diversity of terrestrial life was assembled over geological time[1-7] and the macroevolutionary importance of processes observed in ecological communities[6,8]. In particular, studies of the fossil record have been central to debates about models of diversification and community dynamics on geological timescales[8], but nominally global and regional-scale patterns have been used to argue for both an "expansionist" diversification paradigm[1,3,4], and for constrained or diversity-dependent diversification[2,7,9-11], resulting in great uncertainty.

We take a different approach to most previous studies by examining how species richness within tetrapod communities has changed through geological time. Assemblages from individual fossil localities represent communities of potentially-interacting species[12], and long-term patterns of change in richness within these communities (=local richness or alpha diversity) provide a strong test of the processes that might regulate diversification at macro-scales[8,12,13]. Early work by Bambach[12] recognised that comparisons of within-community richness offer a way to at least partially circumvent many of the sampling biases that confound regional and global diversity curves derived from fossil data[7,10]. Our analysis specifically asks how the richness of tetrapod communities has changed through geological time. Patterns of local richness are related to global and regional patterns through spatial turnover and nestedness (beta diversity and provinciality) [12,14]. Addressing how richness has changed at these larger scales requires different forms of analysis (e.g. [15,16]) that are not utilised here. Nevertheless, our findings regarding local richness are relevant to debates about the processes that could stabilise diversification, either by limiting increases to species counts in local communities[6,8,14], or by allowing them to remain high in spite of decreasing global richness[17,18].

The expansionist model represents unconstrained diversification. It predicts continuous, and typically also large[3], increases in local richness over geologically long intervals, spanning tens of millions of years. Within this paradigm, mass extinctions result in temporary setbacks to an otherwise upward diversity trajectory (e.g. [3]). In contrast, constrained models of diversification predict high rates of origination (speciation plus immigration) whenever lineages can exploit unoccupied ecospace. These abrupt increases in local richness are followed by extended intervals of relative stasis, as origination and extinction rates equilibrate to zero net diversification with increasing diversity[19]. Such slowdowns could be caused by negative biotic interactions within communities, such as

competition for finite resources[14]. Equilibria may also be reset episodically by ecological disruptions such as mass extinctions, major environmental changes, or the evolution of key innovations[8,20]. Patterns of local richness consistent with constrained diversification have been documented for several fossil groups[21,22]. However, local richness in fossil tetrapods is remarkably understudied and prior work is limited in its taxonomic, temporal and/or geographic scope[2,7,23,24].

We document patterns of local richness in non-flying, non-marine (=terrestrial) tetrapods through their entire Phanerozoic history. We also dissect patterns for key subtaxa, including non-avian dinosaurs, mammals and squamates. Flying tetrapods were analyzed separately, because their fragile skeletons result in a much more unevenly sampled fossil record[25]. Our data, drawn from the Paleobiology Database[26], comprise 145,332 taxon occurrences from 27,531 collections. We estimate local richness by counting species and genera per collection, including those specifically-indeterminate occurrences that must represent distinct species, because they record the presence of a higher taxon that is otherwise unknown in the collection (see Methods). Counts of taxa per collection are a widely-accepted proxy for local richness in the fossil record[12]. We also present counts of taxa per geological formation, which broadly corresponds to landscape-scale richness[23].

We use two main forms of analysis to test hypotheses about the processes governing change in local richness through geological time: 1) linear model comparisons representing phases of “expansionist” and “constrained” diversification, using information only from exceptional localities that were parsed according to a set of objective, numerical criteria; and 2) simulated null distributions based on resampling (with replacement) of empirical counts of species from the full set of localities, pooled according to time intervals. Both suggest a similar pattern comprising extended phases of stasis lasting tens of millions of years, interrupted by geologically-abrupt phase-shifts bringing about large increases in richness.

Results and discussion

Visual appraisal suggests that tetrapod local richness (counts of species per collection) experienced long periods of stability during the Permian–Triassic and from either the latest Cretaceous or earliest Cenozoic to the Recent, interrupted by an abrupt increase (in geological terms—spanning hundreds of thousands to, at most, a few million years; Fig. 1a). Total increases in local species richness were small compared to those implied by previous expansionist interpretations (entailing order-of-magnitude increases over the last 100 million years[1,3])—compared to the richest Palaeozoic localities, local species richness had increased at most two-fold by the mid-Mesozoic, and at most three-fold by the latest Cretaceous–early Cenozoic. Following an initial slow increase in the early to mid-Carboniferous, observed local species richness rose steeply in the late Carboniferous (~25 species). Levels in the Permian did not greatly exceed those of the late Carboniferous (~30 species) and remained similar during the Late Triassic, 100 million years later, despite considerably more intense sampling. Patterns of local genus richness are similar (Supplementary Fig. 1), whereas landscape-scale richness (counts of species and genera per formation; Supplementary Fig. 2) suggest a more prolonged initial rise, lasting until the end of the Permian. The observed richness of exceptional collections (species and genera) increased by up to ~2.5 times between the end of the Triassic and the latest Cretaceous. Levels of local richness exceeding those from all earlier intervals occur in the Kimmeridgian–Tithonian stages of the Late Jurassic (~55 species; approximately double that of the Permian–Triassic), and in the Maastrichtian stage of the latest Cretaceous (~70–80 species; around 1.5 times that of the Late Jurassic). High observed local richness in the Kimmeridgian–Tithonian and Maastrichtian is driven by exceptional sampling of small-bodied taxa (mammals, squamates, turtles and lissamphibians). However, per-formation richness of non-flying terrestrial tetrapods at all taxonomic levels suggests little if any Mesozoic increase until the Campanian

(with the exception of the geographically-vast Morrison Formation, Late Jurassic, USA; Supplementary Fig. 2).

Simulated null distributions suggest that much (but not all) of the apparent variation in maximal local richness among intervals results from variation in sampling intensity. These support an interpretation of stasis in local richness interrupted by a single phase-shift following the end-Cretaceous extinction. Empirical patterns are substantially different to those obtained if richness per collection is randomly sampled from a single Phanerozoic-long pool. However, they deviate little from those obtained by random sampling from two pooled distributions divided by the Cretaceous/Palaeogene (K/Pg) boundary: one representing the Carboniferous–Cretaceous and the other representing the Cenozoic (Figs 2 and 3; Supplementary Figs 3 and 4; see Supplementary Information for full description of simulated null distributions). Adding an additional underlying pool to partition the data by geological era only marginally improves the fit over the two-phase pre-/post-K/Pg model, and is therefore not justified (Fig. 3). Although local richness clearly must have increased during the initial Palaeozoic radiation of terrestrial tetrapods, sampling is too limited to confidently resolve early changes in richness through comparison with the simulated null distributions (confidence intervals for the null are very large).

The weight of evidence supports a substantial increase in local richness across the K/Pg, followed by relative stasis toward the present (Fig. 1). A pattern of stasis in local richness before and after the K/Pg, broken by an abrupt two- to three-fold increase, is supported by both rarefaction curves of terrestrial tetrapod local richness quantiles for period-level bins (the Cretaceous is similar in richness to the Jurassic and Triassic but lower than the Paleogene, Neogene or Quaternary; Fig. 4; Supplementary Fig. 5), and comparisons between empirical curves of local richness quantiles and simulated null distributions (Figs 2 and 3). This is consistent with patterns of continent-scale sampling-standardized species richness[10]

and counts of species per formation. This interpretation might appear to be contradicted by the fact that the two richest Maastrichtian localities (Bushy Tailed Blowout and Lull 2 Quarry from the Lance Formation, USA) have species counts comparable to the richest Cenozoic localities. However, the K/Pg boundary has been subject to extraordinarily intense study, which could dramatically inflate the maximum observed estimates of local richness relative to earlier time intervals. These two localities are also substantially richer than all other Maastrichtian sites. On balance, we consider it more likely that tetrapod local richness increased immediately after the K/Pg boundary, rather than just before it.

Linear model comparisons likewise favour constrained models. The "species ~ phase" model explains richness of exceptional localities using only one covariate, describing the existence of distinct phases (see Methods) that are characterised by different average levels of local richness, and otherwise implies no continuous change in richness through time. This model is favoured for five of the nine numerically-defined sets of exceptional localities (median Akaike weight of 0.68, ranging from 0.051–0.74; median adjusted R-squared of 0.679, ranging from 0.511–0.766), although the "species ~ time + phase" model is sometimes favoured when lower richness quantile thresholds are used ("species ~ time + phase"; median Akaike weight of 0.276, ranging from 0.232–0.712; median adjusted R-squared of 0.679, ranging from 0.519–0.762; Supplementary Fig. 6; Supplementary Table 1). Models where time is the only explanatory variable, representing the expansionist paradigm, receive negligible support (median Akaike weight near zero). Furthermore, interval-specific regressions of richnesses of exceptional localities against time are non-significant, with near-zero slopes, for the late Carboniferous–Triassic (spanning ~160 million years) and Cenozoic (66 million years; see Supplementary Information).

Major groups of tetrapods also individually show early or stepwise increases, followed by extended periods of stasis. Observed local richness of non-avian dinosaurs

within exceptional localities rose steadily from the Late Triassic, reaching an apparent maximum in the Late Jurassic, followed by a slightly lower peak in the Campanian–Maastrichtian (Fig. 5a). Rarefaction curves of richness quantiles show that although non-avian dinosaur local richness was clearly lower in the Triassic than in the later Mesozoic, confidence intervals for the Jurassic and Cretaceous strongly overlap (Fig. 6a). Apparent peaks in Mesozoic mammals mirror those for dinosaurs (Fig. 5b), and confidence intervals for richness-quantile rarefaction curves also strongly overlap between the Jurassic and Cretaceous (Fig. 6b). Therefore, the maximum observed levels of local richness for two ecologically-important groups of large- and small-bodied tetrapods did not increase during an interval of nearly 100 million years between the Late Jurassic and end-Cretaceous. However, mammalian local richness increased abruptly by two- to three-fold across the K/Pg boundary (consistent with past studies documenting an extremely rapid recovery after the end-Cretaceous mass extinction[27,28]), reaching a second equilibrium that was maintained throughout the Cenozoic. It is plausible that the magnitude of apparent increase across the K/Pg could be exaggerated by the greater diagnosability of mammal teeth, but the contribution of this effect, if any, is unknown at present (see Methods). Magnitudes of local richness for Cenozoic mammals (60–70 species per collection or 100–200 species per formation) overlap with those obtained for surveys of present-day local faunas in Kenya (60–107 species for landscape-scale survey sites (120–1510 km^2 in area) [29]. Squamate local richness changed little from their first appearance in the Middle Jurassic until the Santonian, but nearly doubled in the Campanian–Maastrichtian, reaching levels that were sustained through most of the Cenozoic (Fig. 5d; Fig. 6c). Further investigation is needed to determine why the recovery of tetrapod communities after the end-Cretaceous mass extinction involved a rapid rebound to higher diversity equilibrium, while that following the end-Permian mass extinction did not.

Importantly, major taphonomic and collector biases show long-term trends of increase (Figs 1, b and c, and Supplementary Fig. 7), reflecting the widely-recognized tendency for geological sampling to improve towards the present[30,31]. Improved sampling towards the present systematically increases the chance of discovering highly diverse localities in younger deposits: better-sampled intervals tend to yield higher local richness estimates (as demonstrated by our simulated null distributions, richness-quantile rarefaction curves, and analyses of correlations between sampling proxies and local richness estimates; see Methods and Supplementary Information; Supplementary Figs 8 and 9). This makes our finding of statistical support for predominantly equilibrial patterns in uncorrected richness data particularly conservative. Correcting for poorer sampling within older deposits using comprehensive abundance data would most likely diminish the observed trend of increasing local richness through time.

The static patterns of local richness we document imply persistent constraints on increases in species counts within communities. Nevertheless, this does not necessarily indicate constraints at regional or global-scale richness, because patterns of richness at different spatial scales may become decoupled via changes in beta diversity or faunal provinciality[14]. Therefore, our results do not entirely preclude a postulated ten-fold expansion in global species richness over the last 100 million years[32]. However, they do demand substantial, and as-yet undocumented, increases in beta diversity (geographic turnover of community composition, or faunal provinciality) to accommodate such a profound increase if it occurred, and this scenario is strongly contradicted by regional-scale studies of sampling-standardised diversity patterns in Mesozoic–Cenozoic terrestrial tetrapods [e.g. [2,7,10]].

Ecologists increasingly regard modern communities as unsaturated[6,33]. A common interpretation of this hypothesis is that local and regional richness should rise in an essentially continuous fashion through palaeontological time[3,6]. This prediction is

inconsistent with a growing portfolio of evidence from the fossil record[7,20,21,23]. We find little support for the expansionist model of diversification within terrestrial tetrapods at the local scale. Instead, local richness shows constrained or asymptotic patterns over hundreds of millions of years, punctuated by rare but abrupt increases during the rise of major groups and following mass extinctions. Importantly, the effects of mass extinctions are very short-lived, with diversity rapidly rebounding to equal or surpass pre-extinction levels.

Whether or not communities are ever truly 'saturated', origination and extinction rates in local assemblages are essentially balanced on timescales of tens of millions of years, leading to equilibrial or constrained diversity patterns. Process-based explanations of this observation are not well-understood[8], and many of the lines of evidence from ecological studies used to support unconstrained diversity dynamics might be compromised if present-day community structures are anomalous for the Phanerozoic due to human impacts[34]. However, our results do not demand that specific kinds of biotic interactions are responsible for observed patterns[35], or even that communities are in equilibrium[36,37]. Reconciling patterns in the fossil record with ecological theory based on modern communities remains a major challenge in evolutionary biology, and one that can only be addressed by integrative studies that unite observations across a wide range of spatiotemporal scales.

Methods

Data download and processing. We downloaded occurrence data for Tetrapodomorpha from the Paleobiology Database[26] (PaleoDB; http://www.paleobiodb.org) on 13 December 2017. The dataset comprises 157,847 taxon occurrences from 33,346 collections (= localities or local faunas) prior to removal of unsuitable data, and 145,332 taxon occurrences from 27,531 collections after cleaning.

We removed occurrences pertaining to marine tetrapods, traces, egg taxa [using lists of names modified from Benson et al. [7]], common wastebasket taxa (e.g. "*Crocodylus*", "*Crocodilus*", "*Alligator*" and "*Lacerta*"), and other dubious occurrences (e.g. non-avian dinosaurs in Cenozoic horizons). Any remaining Cenozoic marine tetrapod occurrences were excluded using sets of occurrence numbers contained within PaleoDB occurrence data downloads for marine mammal clades (Cetacea, Pinnipedimorpha and Sirenia; we did not remove any transitional forms that may lie outside these groups). Marine birds and snakes were not excluded, but make up relatively few occurrences. Occurrences for which preservation mode was listed as "trace", "cast,trace" or "mold/impression,trace" were also removed. Collections with soft-tissue preservation (e.g. Lagerstätten deposits) were retained. Obvious "wastebasket" collections that formed distinct outliers (shown on Supplementary Fig.10; e.g. PaleoDB collection no. 13779, a collection from the Eocene locality of Gran Barranca, Argentina, that was created to house historical specimens not clearly linked in the literature to specific fossil sites) were excluded.

We removed collections with the largest geographic scale of "basin" or the largest stratigraphic scale of "group". This is because changes in the scope of the geographic and temporal sampling universes may bias estimates of local richness. An idealized PaleoDB fossil collection represents an assemblage of fossils originating from a single stratigraphic horizon (i.e., a single bed, or group of beds) within a small geographic area (e.g. a quarry). However, due to differential reporting of stratigraphic and geographic data in the literature, collections may in practice represent geographic scales ranging from single hand-samples, to single outcrops, groups of outcrops, local (such as a series of outcrops occurring over several kilometers), and even basinal (e.g. fossils reported from coming from a particular stratigraphic unit, but without any detailed locality information provided) areas, and stratigraphic scales ranging from a single bed or group of beds to member, formation or

group-level scales. We retained “formation”-level collections, because PaleoDB enterers sometimes assign “formation” level stratigraphic scale even though the collection in question in fact occupies a single bed, simply because more precise stratigraphic information is not given in the literature.

To analyze individual clades, we taxonomically filtered the tetrapod occurrence dataset using occurrence ID numbers contained within separate PaleoDB data downloads for non-avian dinosaurs (Dinosauria excluding Aves), birds (Aves), non-flying mammals (Mammaliaformes excluding Chiroptera), bats (Chiroptera), squamates (Squamata), turtles (Testudinata), and crocodylians and their stem-group (Pseudosuchia).

Time bins. Composite time bins of approximately equal length were used to calculate simulated null distributions, trends in sampling biases and correlations between sampling variables (see below). This binning scheme was based on that used by Benson et al. [7], which was in turn modified from the scheme used by the original Paleobiology Database portal (now Fossilworks; http://www.fossilworks.org). On average, these “equal-length” bins are ~9 myr in duration, but range from 19.5 myr for Tr4 (Norian) to 1.806 for Ng4 (Calabrian, Middle Pleistocene, Late Pleistocene and Holocene). Definitions of equal-length bins are given in Supplementary Table 2. Data were binned by midpoint ages of collections or formations.

Estimating local richness. Fossil localities comprise assemblages of potentially interacting species but, in contrast to instantaneous snapshots from present-day communities, assemblages are generally time-averaged on scales of tens to hundreds of millennia. Measurements of local richness over palaeontological timescales are important because they smooth out short-term fluctuations due to nonequilibrial processes[8] that may obscure longer-term patterns.

We estimate local richness using simple face-value counts of taxa present within localities because sampling-standardised estimates of richness within collections require abundance data, which are not systematically available. Nevertheless, our use of face-value counts allows direct comparison with older studies that used face-value global richness counts to argue for expansionist diversification[1,3]. All of our counts of species include specifically-determinate occurrences plus "indeterminates" — occurrences that are not resolved to species level, but nevertheless indicate the presence of a distinct species in that collection. For example, an occurrence of "Muridae indet." would be counted only if there were no occurrences of murids resolved to finer taxonomic levels within the collection. This measure provides the most accurate estimate of species-level richness, because it tallies every distinct species recognized in a collection regardless of whether or not it has been identified to species level.

To determine whether specifically-indeterminate occurrences represented distinct species, we used the hierarchy of taxonomic names in the PaleoDB. This was achieved by: 1) determining all of the unique accepted names represented by occurrences within a collection; 2) obtaining all of the names in the taxonomic hierarchy above each of these accepted names; and 3) finding which accepted names were not present in the pooled list of unique names drawn from the taxonomic hierarchies. Informal species identified within a collection (such as "Chiroptera informal indet. sp. 1") were counted as separate species. For full operational details, see our R function 'countLocalRichness' in the supplementary analysis files. Counts of genus richness likewise include generically-indeterminate occurrences that must represent distinct genera, using the same procedure outlined above (although without tallying "informal indet." species).

Patterns of local richness were visualized by plotting counts of taxa per collection or formation against their midpoint ages. Although we focus on "global" patterns from

aggregated global data, we also plotted regional patterns of local richness for five continents (North America, South America, Europe, Africa and Asia; note that Russia was assigned to Asia). Patterns of local richness were also dissected by palaeolatitudinal zone (low = 0°–30°, mid = 30°–60°, and high = 60°–90° palaeolatitude).

Defining exceptional localities using objective numerical criteria. Unlike extant ecosystems, fossil samples—especially for terrestrial tetrapods—are characterised by strongly negatively-skewed distributions of richness per collection. This occurs because of a pervasive issue of sample incompleteness: most localities only record a tiny fraction of the original biota, and are thus largely uninformative about patterns of local richness. We therefore focus our qualitative interpretations of the data, and our quantitative linear model comparisons, on exceptional, extensively-sampled localities. These record the most complete snapshots of ancient ecosystems, although they are relatively few in number (Fig. 1a, Supplementary Figs 6 and 11; key information about a selected set of exceptional localities is given in Supplementary Table 3).

We identify these using objective numerical criteria based upon co-occurrences of higher taxa (e.g., including both small-bodied groups such as mammals, and large-bodied groups such as dinosaurs) and richness quantiles. Richness quantiles were calculated within period-level bins. We explored parameter space using nine objective sets of exceptional localities, created by combining three quantile thresholds for non-flying terrestrial tetrapod species richness (0.99, 0.995 and 0.998) and three levels of taxonomic co-occurrence criteria ("None" = no restrictions; "Moderate" = localities containing at least one mammal, one dinosaur and one squamate during their ranges; and "Stringent" = at least one mammal, one dinosaur, one squamate, one turtle, and either a pseudosuchian, lissamphibian or flying tetrapod). These taxonomic co-occurrence criteria ensure that both large- and small-bodied taxa are present. Because most of these major groups did not arise until later in the Mesozoic,

prior to 200 Ma localities could be defined as exceptional only if they contained at least 20 species (“Moderate” and “Stringent” levels of taxonomic co-occurrence criteria only).

Quantitative analyses of local richness. To test alternative hypotheses about modes of diversification (e.g., evaluating the relative support for continuous/gradual/exponential increases through time versus periods of stasis broken by sudden changes in diversity equilibria), we fitted a range of linear models to values of non-flying terrestrial tetrapod local richness within exceptional localities. These linear models represent either: 1) continual increases in richness through time (expansionist models), or 2) stepwise increases separated by extending intervals of approximate stasis, during which short-term variability does not accumulate into longer-term increases (constrained models).

We also document trends in key biases that influence the preservation and recovery potential of fossils, and therefore inform our interpretations. We document substantial increases through the latest Cretaceous and Cenozoic in: 1) the number of fossil collections and their geographic spread; 2) the frequency and number of bulk-sampled collections; 3) collections derived from unlithified or poorly-lithified sediments; 4) collections from low paleolatitudes (Supplementary Fig. 7); and 5) collections from depositional environments that do not preserve well in deep time (see discussion below; Supplementary Fig. 7c). All of these lines of evidence indicate that younger deposits either favor the preservation of rich faunas or allow easier extraction and diagnosis of fossil specimens. This progressively expands the size of the accessible taxonomic sampling universe, driving increases in within-collection sample completeness, and elevating raw estimates of local richness nearer the present.

We controlled for sampling-intensity biases on inferred patterns of local richness using three approaches: by simulating null distributions of local richness quantiles; generating rarefaction curves of per-bin local richness; and analysing correlations between sampling proxies and per-bin local richness estimates. Together, these analyses suggest that:

1) support for our interpretation is substantially strengthened by controlling for variation in sampling intensity through time; 2) our interpretation is robust to the use of different subsets of the data; and 3) variation in factors such as environments are relatively unimportant for the types of analyses done here, having only small effects on inferred patterns.

*Testing diversification hypotheses***.** We used the Akaike Information Criterion with small-sample-size correction (AICc) to evaluate the relative fit of an intercept-only null model (using the formula "species ~ 1") of static non-flying terrestrial tetrapod richness against a simple linear model of richness as a function of time, plus multiple regressions incorporating time and/or a "diversification phase" factor as a covariate or interaction term[38,39]. The phase model uses two covariates: time, and a categorical variable consisting of temporal intervals corresponding to diversification phases (phase1 = Devonian–Mississippian, phase 2 = Pennsylvanian–Triassic, phase 3 = Jurassic–Cretaceous [beginning in the Tr4 time bin for regressions on individual phases, because the Early Jurassic lacks exceptional localities] and phase 4 = Maastrichtian–Cenozoic), allowing the intercept to vary through time while sharing the same slope. No exceptional localities are known from Phase 1. Additionally, we fitted individual regressions to the diversification phases defined above. We did not attempt to fit a multiphase logistic model because the temporal resolution of the data is too coarse and the density of data points too low in many key intervals, particularly in the early Carboniferous and Early to Middle Jurassic.

Sampling intensity biases. Opportunities for sampling terrestrial localities increase dramatically nearer the present, driven by the increasing availability of fossil-bearing sediments from the Late Cretaceous onwards: over half of all the exposed rocks from terrestrial environments date from the Cretaceous and Miocene[30]. There are substantial increases in per-bin counts of collections and occupied equal-area grid cells (a measure of palaeogeographic spread; Fig. 1c [variables log-transformed] and Supplementary Fig. 7 [un-

transformed variables]), especially from low palaeolatitudes, which are poorly known during earlier intervals.

Unlike global or regional richness, increased sampling of localities or geographic regions does not directly or mechanistically cause local richness to increase. Estimates of local richness are not additive with respect to sampling of new localities—rather, they are drawn probabilistically from an underlying distribution. The process is analogous to playing a slot machine: a single attempt may yield a spectacular win, while many attempts may fail to net a return. All else being equal, however, playing more frequently (i.e., sampling the fossil record more intensively) increases one's chances of winning. As a result, it is important to analytically control for sampling intensity when documenting local richness patterns; we do this using simulated null distributions, rarefaction curves of local richness quantiles and correlation tests.

Simulated null distributions. Our simulated null distributions make use of a different principle to the linear models. Instead of focusing only on exceptional localities, they simulated how variation in counts of localities through time could bias curves of per-bin richness quantiles when values of richness per collection are drawn probabilistically from a fixed underlying distribution. This distribution is simply the full set of per-collection richness values, and is either 1) lumped to a single, Phanerozoic-long pool, or 2) split into larger numbers of smaller pools delimited by time, and represent hypotheses about phase-shifts in local richness equilibria.

The simulations comprised 1000 independent trials. In each trial, the richness value for each collection in the full Phanerozoic dataset was drawn at random, with replacement, from a pooled set of local richness values. Four pools were used: a single Phanerozoic pool; pre- and post-K/Pg pools, era pools (Pz, Mz and Cz), and diversification-phase pools (with boundaries at the T/J and K/Pg boundaries). Simulations using multiple pools represent

hypotheses about shifts in local richness. Richness quantiles were then calculated for each equal-length bin (0.5, 0.75, 0.9, 0.95, 0.99 and 1 [1 = the per-bin maximum]). Mean values for per-bin richness quantiles were then calculated across all trials, along with 95% CIs. Confidence intervals for the empirical curves were calculated using 999 bootstrap replicates. Both null and empirical curves were smoothed using loess regression to de-emphasise bin-to-bin noise, but the null distributions use a slightly higher 'span' to increase smoothing (0.25 vs 0.1 for the empirical curve).

Rarefaction curves of local richness quantiles. We also used subsampling to generate rarefaction curves for geological periods, allowing per-bin quantiles of local richness for terrestrial tetrapods and major subgroups to be compared at different levels of comparable sampling intensity. We calculated rarefaction curves of local richness quantiles (at quantile levels ranging from 0.9–1), rarefied by collection, within period-length bins. Richness quantiles below 0.9 were found to be uninformative, as they were too heavily leveraged by abundant but highly depauperate localities. These curves reveal the expected local richness quantile values across repeated draws of a fixed number of collections, thus controlling for the number of sampling opportunities.

Correlation analyses. Lastly, correlation analyses show how per-bin changes in local richness relate to sampling intensity on long and short timescales. It is important to assess the short-term correlation between local richness and sampling proxies, because this allows us to evaluate the degree to which sampling effort directly controls within-collection richness. To determine the relative influence of short- and long-term factors governing the relationship between local richness quantiles and sampling, as quantified by counts of collections and equal-area occupied grid-cells, we evaluated: 1) the correlation between raw variables using Spearman’s rank-order correlation; and 2) the correlation between variables after detrending the data series using an ARIMA model. All variables were log-transformed.

The optimal degree of differencing and values for autoregressive and moving average components in ARIMA models for each data series were automatically chosen using the function auto.arima in the R package forecast[40]. If the correlation after detrending is substantially stronger than the raw correlation, then the relationship must be driven by long-term trends (such as a long-term increase in the intensity and distribution of sampling), rather than short-term fluctuations in sampling.

Lithification biases. It also becomes easier to find and extract fossils from younger deposits. A substantial rise in counts and proportions of unlithified or poorly-lithified terrestrial sediments, accompanied by similar increases in the use of bulk-sampling methods, begins in the Late Cretaceous (Fig. 1c, Supplementary Figs 7b and 12). Bulk-sampling of unlithified sediments permits more exhaustive sampling of a local fauna[41]. This should cause within-site sample completeness, and thus raw estimates of local richness, to increase systematically towards the present[41,42]. For example, the lithification bias might account for as much as half of the three- to four-fold increases in local richness of marine invertebrates during the Cenozoic[43]. Notably, there is a pronounced peak in counts of bulk-sampled localities during the Campanian–Maastrichtian that might account for the apparent end-Cretaceous rise in tetrapod local richness (Supplementary Fig. 7). The lithification bias can be partially addressed by applying sampling-standardization methods. However, the abundance data required are not generally available in the literature for Phanerozoic tetrapods.

Environmental heterogeneity. We cannot directly account for variation in the environments represented by exceptional localities through time but we can show that, with one important exception, there are not substantial systematic changes in these environments through the study interval (athough there is usually considerable variation; see Supplementary Figs 13 and 14; see also Supplementary Table 3). The one exception results

from the large increases in geologically-ephemeral environments such as cave deposits within the past five to ten million years[31], which we excluded from our analyses (Supplementary Fig. 7c). In the Plio–Pleistocene, lithification biases combine with the wide geographic distribution of depositional environments that are poorly represented in the deeper fossil record to markedly increase the quality of sampling at the local scale.

These environments, dubbed "doomed sediments" by Holland[31], include cave-fill deposits, tar-pits and many fluvial and lacustrine environments not located in subsiding basins and which only rarely survive into the deep-time rock record because they are more frequently erased by early erosion. Such environments are not definitively "doomed" over long timescales, but they are substantially more prevalent near the present than in deep time. The progressive loss of deposits from these depositional environments creates an increasing preservational discrepancy over time and systematically changes the rate of fossil preservation from shallow to deep time[31]. In particular, the large increase in Plio-Pleistocene cave-fill environments, which preserve groups that are rarely fossilized, such as birds, bats and lissamphibians[44], drives a coincident spike in sampled local richness of flying tetrapods (Supplementary Fig. 15). We therefore omit Plio–Pleistocene karst environments from our main figures and analyses. However, ancient cave-systems pre-dating the Plio-Pleistocene (identified in the PaleoDB as general "karst" or "fissure-fill" environments) were retained. These karst deposits, especially fissure-fills and sink holes, provide us with rare but exceptional and important windows onto ancient faunas, such as the Triassic/Jurassic fissure fill deposits from the southwest UK and Richards Spur, from Permian deposits in the Midwest USA.

Taphonomic biases. Taphonomic factors causing differences in preservation potential among higher taxa may bias local richness of individual groups up or down. Nevertheless, this should not obscure relative changes through time, so long as these biases are consistent

494 through time. However, consistently high levels of non-flying terrestrial tetrapod local
495 richness during the Cenozoic may be in part an artifact of the ecological ascendancy of
496 crown-group mammals (including increases in richness and abundance). Durable mammal
497 teeth are easily preserved and diagnosed from limited fossil material. This permits more
498 consistently fair comparisons of richness through time, and results in higher apparent local
499 richness than that of groups that were likely at least as diverse. For example, today, on a
500 global level, squamates are more than twice as diverse as non-flying mammals, yet the
501 apparent local richness of mammals is two- to three-fold greater over much of the Cenozoic
502 (Fig. 5, b and c).

503 *Geographic biases*. Shifts in spatial sampling through deep time could potentially
504 affect apparent patterns of local richness. Today, latitudinal climate variation is a key driver
505 of richness, but sampling in our dataset is dominated by records from temperate
506 palaeolatitudes (Supplementary Fig. 16). Patterns from well-sampled continents such as
507 Europe and North America are comparable to those of the aggregated global data
508 (Supplementary Fig.17). However, comparisons between the Palaeozoic and later intervals
509 may be complicated because the record up until the mid-Permian is predominantly
510 palaeoequatorial, whereas much of the later record derives from palaeotemperate regions.
511 This bias could inflate Palaeozoic local richness relative to later time intervals.

512 To investigate the potential influence of preservational factors on patterns of tetrapod
513 local richness, we visualized the distribution of collections representing poorly-lithified or
514 unlithified sediments, those that had been bulk-sampled, and those originating from low
515 palaeolatitudes (within 30° of latitude from the palaeoequator). There are particularly large
516 Cenozoic increases in sampling at low palaeolatitudes[45] (Supplementary Fig. 7c). Large-scale
517 terrestrial rock-record biases are also affected by the proximity to the modern equator, where
518 outcrops are limited by greater vegetation cover[46]. Additionally, rapid decay of organic

material near the terrestrial palaeoequator reduces the chance of fossilization in the first place[47], a bias which would work in opposition to recovering modern-style latitudinal diversity gradients in the fossil record.

Life Sciences Reporting Summary. Further information on experimental design is available in the Life Sciences Reporting Summary.

Data and code availability. The data used in this study were downloaded from the Paleobiology Database (http://www.paleobiodb.org) and have been archived, together with all custom analysis scripts, on Dryad (https://doi.org/10.5061/dryad.3v0p84v).

References

1. Benton, M. J. Diversification and extinction in the history of life. *Science* **268,** 52–58 (1995).
2. Alroy, J. in *Speciation and Patterns of Diversity* (eds. Butlin, R., Bridle, J. & Schluter, D.) 301–323 (Cambridge University Press, 2009). doi:10.1017/CBO9780511815683.017
3. Benton, M. J. & Emerson, B. C. How did life become so diverse? The dynamics of diversification according to the fossil record and molecular phylogenetics. *Palaeontology* **50,** 23–40 (2007).
4. Kalmar, A. & Currie, D. J. The completeness of the continental fossil record and its impact on patterns of diversification. *Paleobiology* **36,** 51–60 (2010).
5. Vermeij, G. J. & Grosberg, R. K. The great divergence: when did diversity on land exceed that in the sea? *Integr. Comp. Biol.* **50,** 675–682 (2010).
6. Harmon, L. J. & Harrison, S. Species diversity is dynamic and unbounded at local and continental scales. *Am. Nat.* **185,** 584–593 (2015).
7. Benson, R. B. J. *et al.* Near-stasis in the long-term diversification of Mesozoic tetrapods. *PLoS Biol.* **14,** e1002359 (2016).
8. Rabosky, D. L. & Hurlbert, A. H. Species richness at continental scales is dominated by ecological limits. *Am. Nat.* **185,** 572–583 (2015).
9. Liow, L. H. & Finarelli, J. A. A dynamic global equilibrium in carnivoran diversification over 20 million years. *Proc. R. Soc. B* **281,** 20132312–20132312 (2014).

10. Close, R. A., Benson, R. B. J., Upchurch, P. & Butler, R. J. Controlling for the species-area effect supports constrained long-term Mesozoic terrestrial vertebrate diversification. *Nat. Commun.* **8,** 15381 (2017).
11. Cantalapiedra, J. L., Domingo, M. S. & Domingo, L. Multi-scale interplays of biotic and abiotic drivers shape mammalian sub-continental diversity over millions of years. *Sci Rep* **8,** 1–8 (2018).
12. Bambach, R. K. Species richness in marine benthic habitats through the Phanerozoic. *Paleobiology* (1977).
13. Wiens, J. J. The causes of species richness patterns across space, time, and clades and the role of 'ecological limits'. *Q. Rev. Biol.* **86,** 75–96 (2011).
14. Alroy, J. Limits to species richness in terrestrial communities. *Ecology* **53,** 1211–9 (2018).
15. Davis, E. B. Mammalian beta diversity in the Great Basin, western USA: palaeontological data suggest deep origin of modern macroecological structure. *Global Ecology and Biogeography* **14,** 479–490 (2005).
16. Vavrek, M. J. & Larsson, H. C. E. Low beta diversity of Maastrichtian dinosaurs of North America. *PNAS* **107,** 8265–8268 (2010).
17. Primack, R. B. *et al.* Biodiversity gains? The debate on changes in local- vs global-scale species richness. *Biological Conservation* **219,** A1–A3 (2018).
18. Cardinale, B. J., Gonzalez, A., Allington, G. R. H. & Loreau, M. Is local biodiversity declining or not? A summary of the debate over analysis of species richness time trends. *Biological Conservation* **219,** 175–183 (2018).
19. Sepkoski, J. J., Jr. A kinetic model of Phanerozoic taxonomic diversity. III. Post-Paleozoic families and mass extinctions. *Paleobiology* **10,** 246–267 (1984).
20. Alroy, J. Dynamics of origination and extinction in the marine fossil record. *PNAS* **105 Suppl 1,** 11536–11542 (2008).
21. Knoll, A. H. in (eds. Diamond, J. & Case, T. J.) 126–141 (Harper & Row, 1986).
22. Powell, M. G. & Kowalewski, M. Increase in evenness and sampled alpha diversity through the Phanerozoic: comparison of early Paleozoic and Cenozoic marine fossil assemblages. *Geology* **30,** 331 (2002).
23. Stucky, R. K. in *Current Mammology* **2,** 375–432 (1990).

24. Barry, J. C. *et al.* Faunal and environmental change in the late Miocene Siwaliks of northern Pakistan. *Paleobiology* **28,** 1–71 (2002).
25. Brocklehurst, N., Upchurch, P., Mannion, P. D. & O'Connor, J. K. The completeness of the fossil record of Mesozoic birds: implications for early avian evolution. *PLoS ONE* **7,** e39056 (2012).
26. The Paleobiology Database.
27. Alroy, J. The Fossil Record of North American Mammals: Evidence for a Paleocene Evolutionary Radiation. *Syst. Biol.* **48,** 107–118 (1999).
28. Wilson, G. P. in *Through the End of the Cretaceous in the Type Locality of the Hell Creek Formation in Montana and Adjacent Areas* **503,** 365–392 (Geological Society of America, 2014).
29. Tóth, A. B., Lyons, S. K. & Behrensmeyer, A. K. A century of change in Kenya's mammal communities: increased richness and decreased uniqueness in six protected areas. *PLoS ONE* **9,** e93092 (2014).
30. Wall, P. D., Ivany, L. C. & Wilkinson, B. H. Impact of outcrop area on estimates of Phanerozoic terrestrial biodiversity trends. *Geol. Soc. London Spec. Publ.* **358,** 53–62 (2011).
31. Holland, S. M. The non-uniformity of fossil preservation. *Philos. T. Roy. Soc. B.* **371,** 20150130 (2016).
32. Benton, M. J. Origins of Biodiversity. *PLoS Biol.* **14,** e2000724 (2016).
33. Mateo, R. G., Mokany, K. & Guisan, A. Biodiversity models: what if unsaturation is the rule? *Trends Ecol. Evol.* (2017). doi:10.1016/j.tree.2017.05.003
34. Lyons, S. K. *et al.* Holocene shifts in the assembly of plant and animal communities implicate human impacts. *Nature* **529,** 80–83 (2016).
35. Hubbell, S. P. *The Unified Neutral Theory of Biodiversity and Biogeography (MPB-32)*. (Princeton University Press, 2011).
36. Connell, J. H. Diversity in tropical rain forests and coral reefs. *Science* **199,** 1302–1310 (1978).
37. Huston, M. A general hypothesis of species diversity. *The American Naturalist* **113,** 81–101 (1979).
38. Benson, R. B. J. & Mannion, P. D. Multi-variate models are essential for understanding vertebrate diversification in deep time. *Biol. Lett.* **8,** 127–130 (2012).
39. Benson, R. B. J. & Upchurch, P. Diversity trends in the establishment of terrestrial vertebrate eco-systems: Interactions between spatial and temporal sampling biases. *Geology* **41,** G33543.1–46 (2012).

40. Hyndman, R. J. & Khandakar, Y. Automatic time series forecasting: the forecast package for R. *J. Stat. Soft.* **27,** (2008).
41. Alroy, J. *et al*. Phanerozoic trends in the global diversity of marine invertebrates. *Science* **321,** 97–100 (2008).
42. Sessa, J. A., Patzkowsky, M. E. & Bralower, T. J. The impact of lithification on the diversity, size distribution, and recovery dynamics of marine invertebrate assemblages. *Geology* **37,** 115–118 (2009).
43. Hendy, A. J. W. The influence of lithification on Cenozoic marine biodiversity trends. *Paleobiology* **35,** 51–62 (2009).
44. Jass, C. N. & George, C. O. An assessment of the contribution of fossil cave deposits to the Quaternary paleontological record. *Quaternary International* **217,** 105–116 (2010).
45. Allison, P. A. & Briggs, D. E. G. Paleolatitudinal Sampling Bias, Phanerozoic Species-Diversity, and the End-Permian Extinction. *Geology* **21,** 65–68 (1993).
46. Noto, C. R. in *Taphonomy* (eds. Allison, P. A. & Bottjer, D. J.) **32,** 287–336 (Springer, 2010).
47. Behrensmeyer, A. K., Kidwell, S. M. & Gastaldo, R. A. Taphonomy and paleobiology. *Paleobiology* **26,** 103–147 (2000).

Supplementary Information is linked to the online version of the paper at http://www.nature.com/nature.

Acknowledgements. We thank all contributors to the Paleobiology Database. This research was funded by the European Union's Horizon 2020 research and innovation programme 2014–2018 under grant agreement 637483 (ERC Starting Grant TERRA, RJB). PDM was supported by a Leverhulme Trust Early Career Fellowship (ECF-2014-662) and a Royal Society University Research Fellowship (UF160216).

Author contributions. RAC, RBJB and RBJ conceived the study. JA, AKB, JB, RBJB, RJB, MTC, TJC, ED, PDM and MU contributed to the dataset. RAC designed and conducted the analyses. RAC wrote the manuscript. RBJB, JA and MTC provided methodological advice. RJB and RBJB drafted portions of the manuscript. All authors provided critical comments on the manuscript.

Competing interests. The authors declare no competing interests.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to RAC (r.a.close@bham.ac.uk).

Figure Captions

Fig. 1. Patterns of local richness for Phanerozoic terrestrial non-flying tetrapods and time series of key fossil-record sampling metrics. a, local species richness (red points denote exceptional localities identified using moderately-strict taxonomic co-occurrence criteria and a richness quantile threshold of 0.995). **b**, Counts of occupied equal-area grid cells (50 km spacing), total counts of collections, counts of bulk-sampled and unlithified or poorly-lithified collections (note logarithmic scale). For age ranges of approximately equal-length bins used in panel b, see Methods.

Fig. 2. Comparison between the empirical curve of local species richness (quantile = 0.9) for terrestrial tetrapods and simulated null distributions. Richness quantiles calculated for equal-length bins. Underlying pools of local richness values for simulated null distributions were drawn either from the whole Phanerozoic, or separate pre-/post-K/Pg or era-level pools (median and 95% confidence intervals calculated from 1000 bootstrap replicates). The null distribution generated from a single Phanerozoic pool of local richness values fits the empirical richness curve poorly. Using separate pre-/post-K/Pg pools substantially improves the fit (see also Fig. 3), while adding a third pool (Era Pools) makes little additional difference. For additional richness quantile levels, see Supplementary Fig. 3.

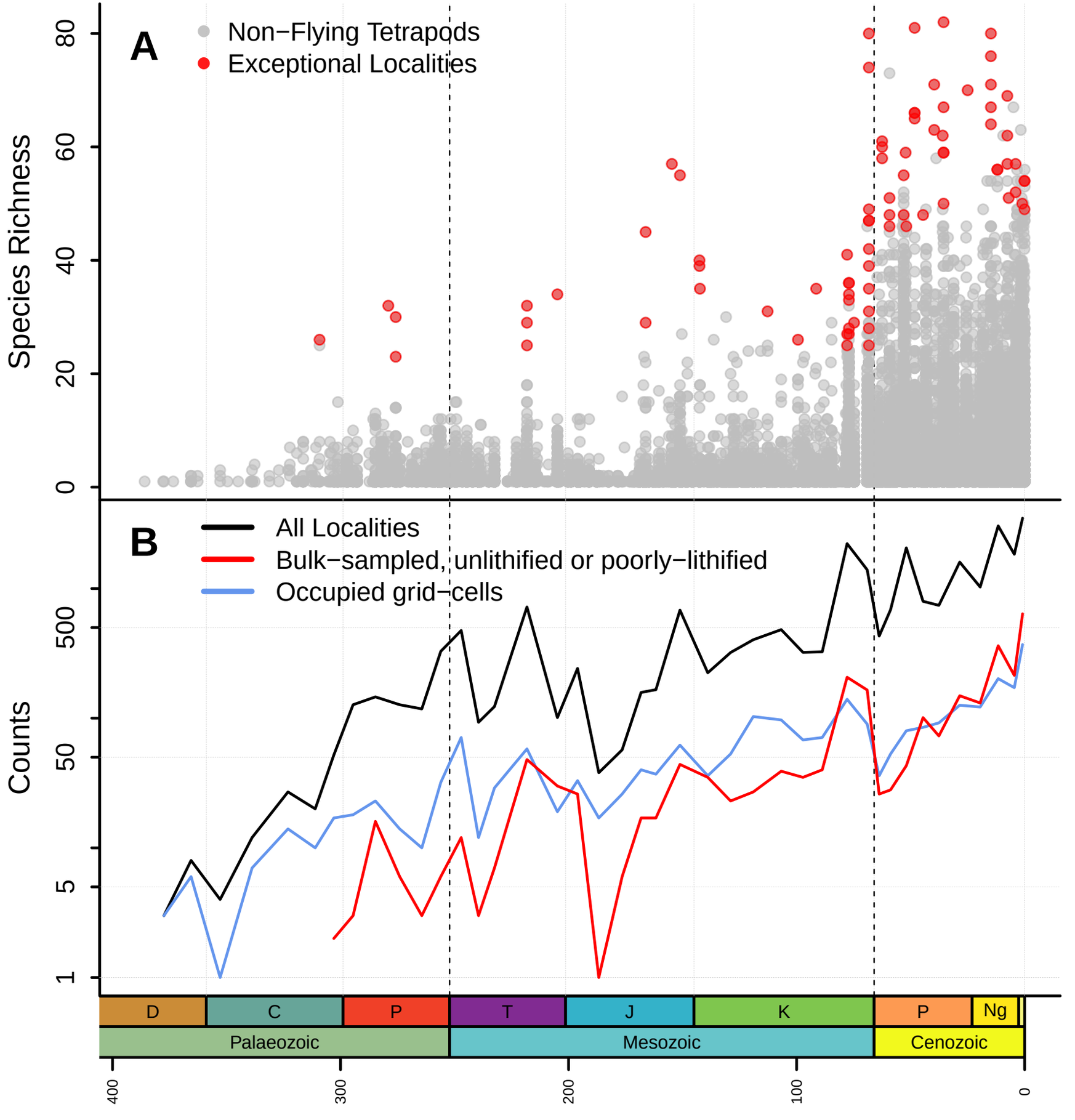
Fig. 3. Residual sums of squares for simulated null distributions of local richness (quantile = 0.9) using Phanerozic, Pre-/Post-K/Pg and Era Pools. Using separate pre-/post-K/Pg pools substantially improves the fit compared to a single Phanerozoic pool, while adding a third pool (Era Pools) makes only a negligible improvement. For additional richness quantile levels, see Supplementary Fig. 4.

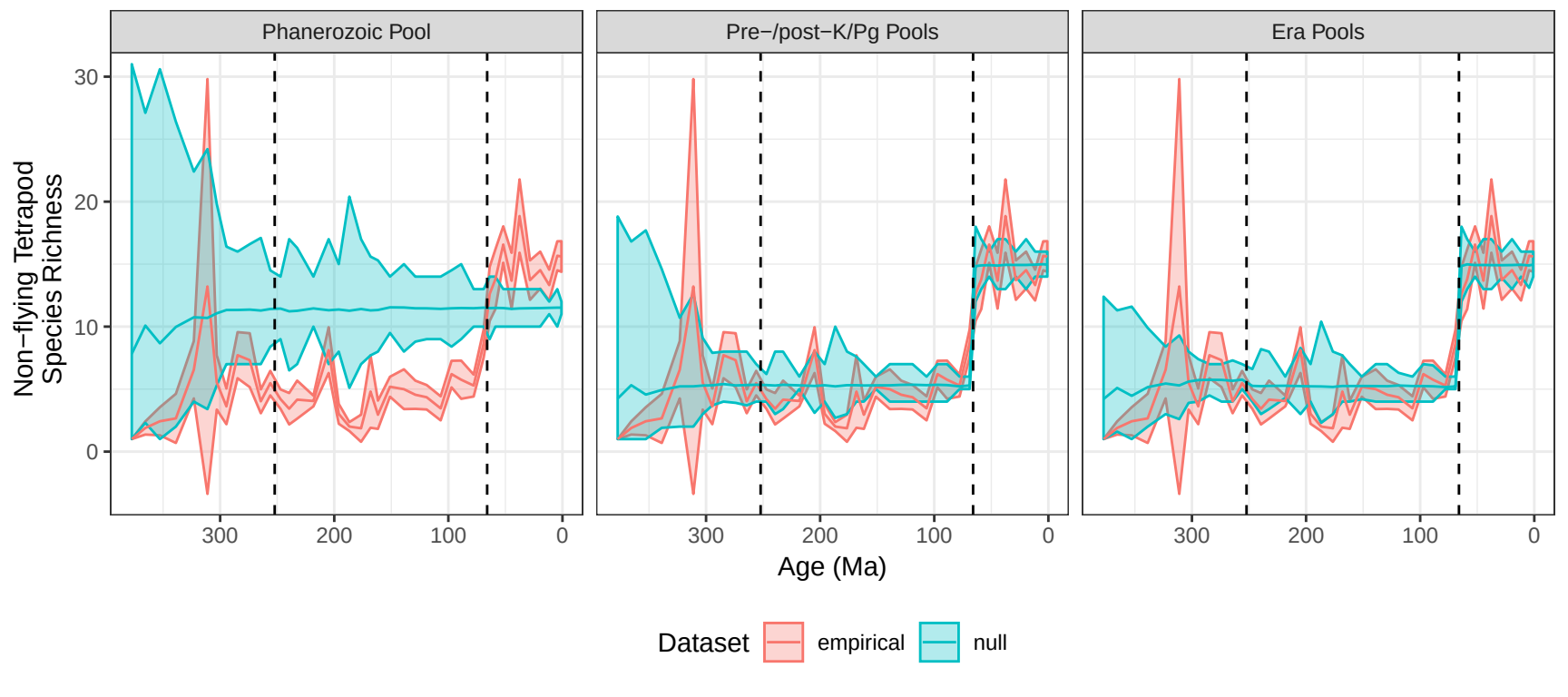
Fig. 4. Rarefaction curves of local richness (quantile = 0.95) per period bin for terrestrial tetrapods. Shaded regions show 95% confidence intervals calculated by

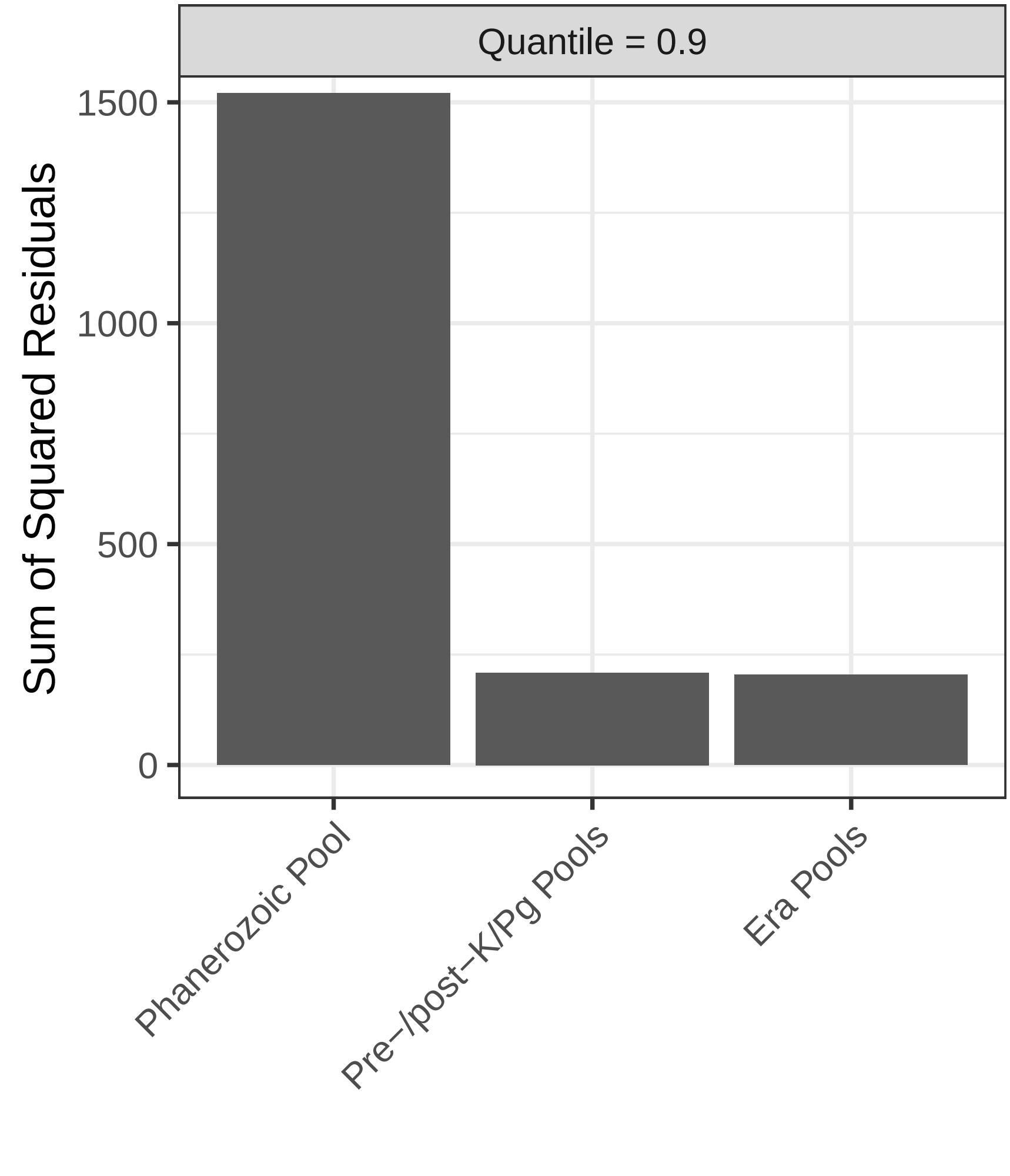
bootstrapping (1000 replicates). **a**, Cenozoic. **b**, Mesozoic. **c**, Palaeozoic. Abbreviations: Q = Quaternary, Ng = Neogene, Pg = Palaeogene, K = Cretaceous, J = Jurassic, T = Triassic, P = Permian, C = Carboniferous, D = Devonian.

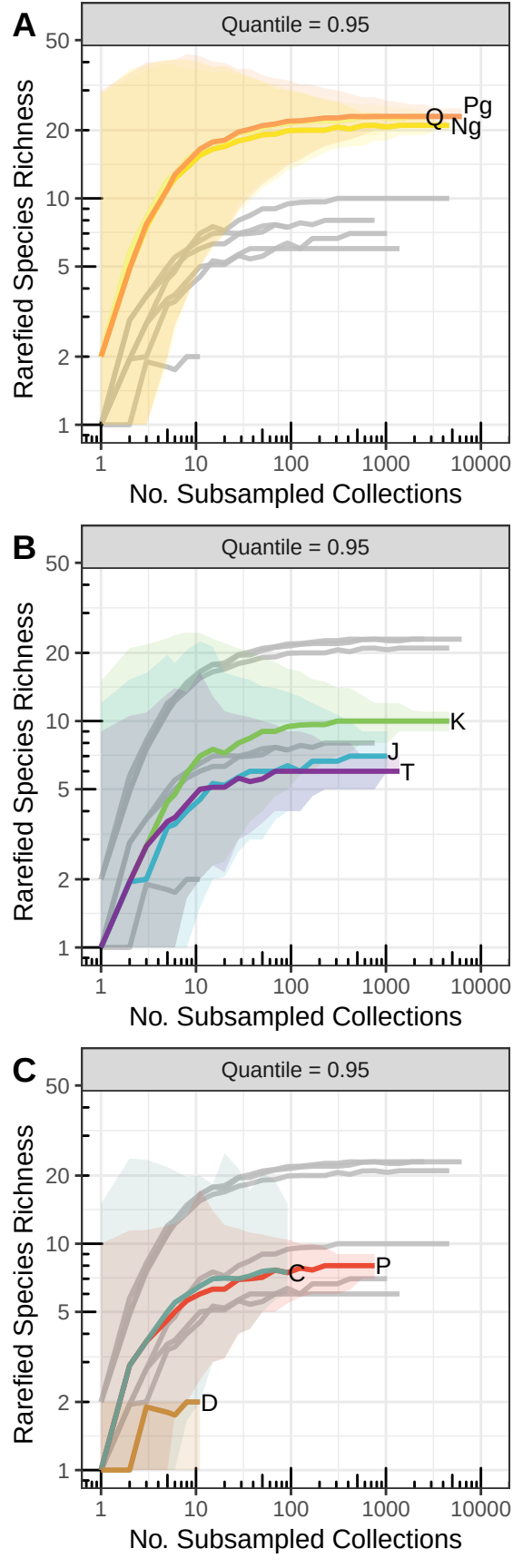
Fig. 5. Clade-level patterns of local species richness. a, non-avian dinosaurs. **b**, non-chiropteran mammaliamorphs. **c**, squamates.

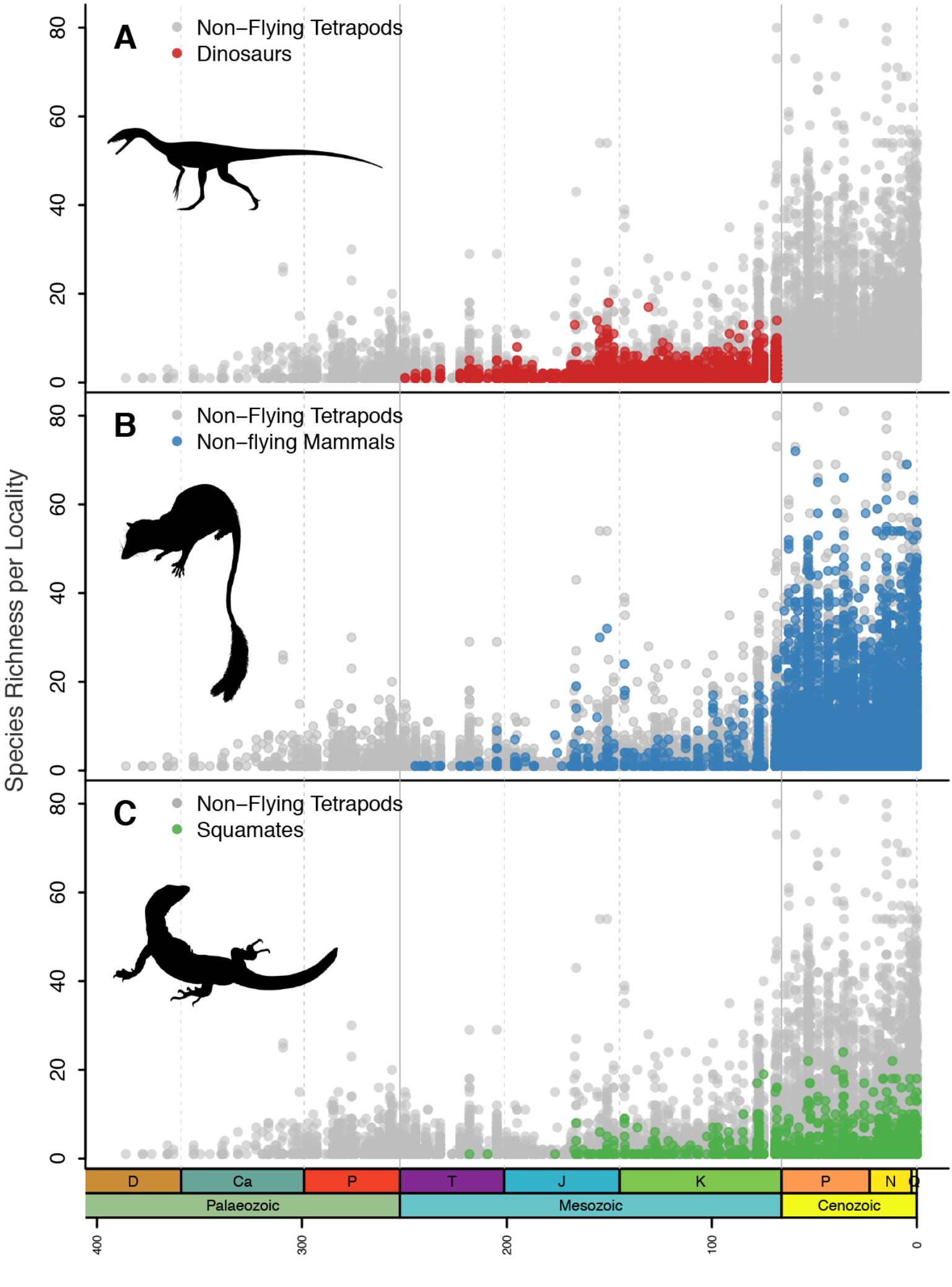
Fig. 6. Rarefaction curves of local richness (quantile = 0.95) per period bin for major tetrapod subclades (non-avian dinosaurs, non-chiropteran mammaliamorphs and squamates). Shaded regions show 95% confidence intervals calculated by bootstrapping (1000 replicates). Abbreviations: Q = Quaternary, Ng = Neogene, Pg = Palaeogene, K = Cretaceous, J = Jurassic, T = Triassic, P = Permian, C = Carboniferous, D = Devonian.

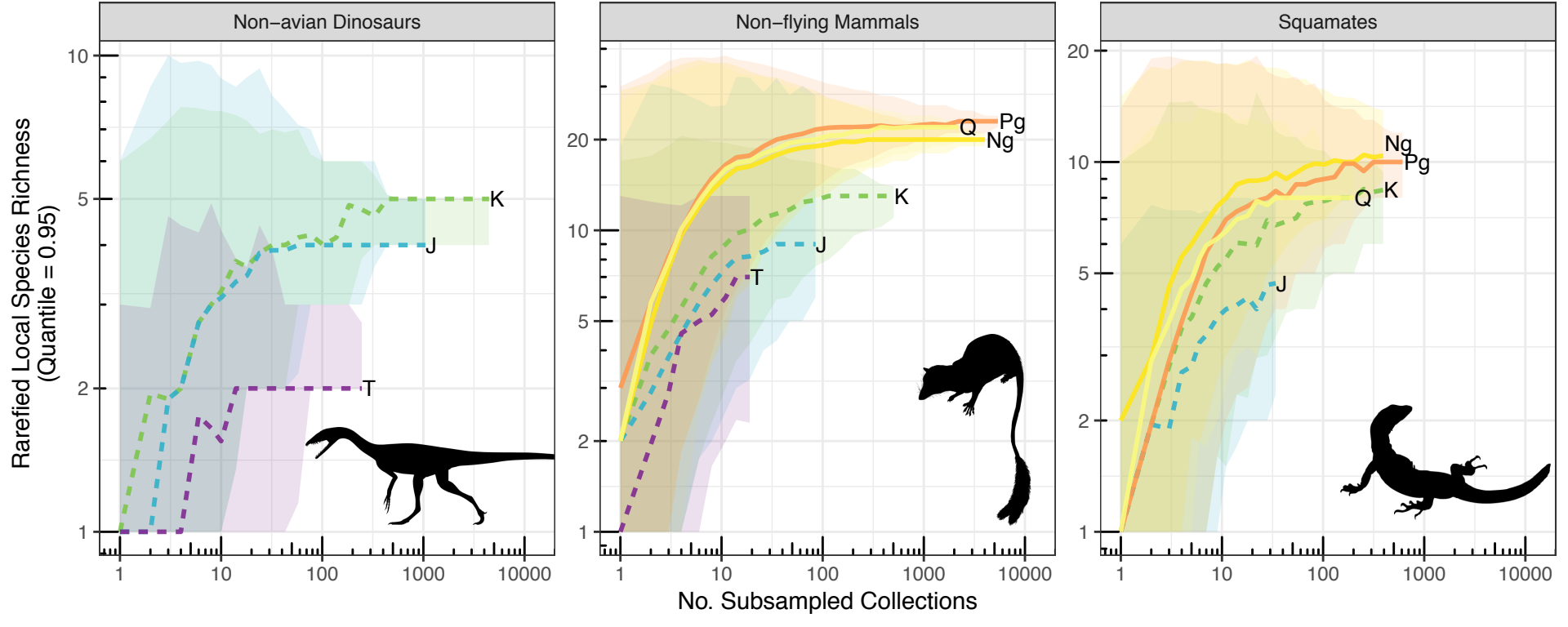

A
Non−Flying Tetrapods
Exceptional Localities
Species Richness
0
20
40
60
80
B
All Localities
Bulk−sampled, unlithified or poorly−lithified
Occupied grid−cells
Counts
1
5
50
500
D
C
P
T
J
K
P
Ng
Palaeozoic
Mesozoic
Cenozoic
400
300
200
100
0


Phanerozoic Pool
Pre−/post−K/Pg Pools
Era Pools
Non−flying Tetrapod Species Richness
30
20
10
0
300
200
100
0
Age (Ma)
Dataset
empirical
null


Quantile = 0.9
Sum of Squared Residuals
1500
1000
500
0
Phanerozoic Pool
Pre−/post−K/Pg Pools
Era Pools


A
Quantile = 0.95
Rarefied Species Richness
50
20
10
5
2
1
Pg
Q
Ng
1
10
100
1000
10000
No. Subsampled Collections
B
Quantile = 0.95
Rarefied Species Richness
50
20
10
5
2
1
K
J
T
1
10
100
1000
10000
No. Subsampled Collections
C
Quantile = 0.95
Rarefied Species Richness
50
20
10
5
2
1
C
P
D
1
10
100
1000
10000
No. Subsampled Collections


A
Non–Flying Tetrapods
Dinosaurs
B
Non–Flying Tetrapods
Non–flying Mammals
C
Non–Flying Tetrapods
Squamates
Species Richness per Locality
0
20
40
60
80
D
Ca
P
T
J
K
P
N
Q
Palaeozoic
Mesozoic
Cenozoic
400
300
200
100
0


Non−avian Dinosaurs
Non−flying Mammals
Squamates
Rarefied Local Species Richness
(Quantile = 0.95)
No. Subsampled Collections
K
J
T
Q
Pg
Ng
1
2
5
10
20
1
10
100
1000
10000